# The Artificial Intelligence and Psychology Project



Departments of
Computer Science and Psychology
Carnegie Mellon University

Learning Research and Development Center
University of Pittsburgh

(2)

# Rationale for a "Many Maps" Phonology Machine

Technical Report AIP - 114


David S. Touretzky, Deirdre W. Wheeler

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Department of Linguistics
University of Pittsburgh
Pittsburgh, PA 15260


March, 1990


This research was supported by the Computer Sciences Division, Office of Naval Research, under contract number N00014-86-K-0678. Reproduction in whole or part is permitted for any purpose of the United States Government. Approved for public release; distribution unlimited.

| | |
|---|---|
| 1a. REPORT SECURITY CLASSIFICATION Unclassified | 1b. RESTRICTIVE MARKINGS |
| 2a. SECURITY CLASSIFICATION AUTHORITY | 3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited |
| 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4. PERFORMING ORGANIZATION REPORT NUMBER(S) AIP - 114 | 5. MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a. NAME OF PERFORMING ORGANIZATION Carnegie Mellon University; 6b. OFFICE SYMBOL (If applicable) | 7a. NAME OF MONITORING ORGANIZATION Computer Sciences Division Office of Naval Research (Code 1133) |
| 6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Pittsburgh, PA 15213 | 7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000 |
| 8a. NAME OF FUNDING/SPONSORING ORGANIZATION Same as Monitoring Organization; 8b. OFFICE SYMBOL (If applicable) | 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0678 |
| 8c. ADDRESS (City, State, and ZIP Code) | 10. SOURCE OF FUNDING NUMBERS p40005ub201/7-4-86 |

| PROGRAM ELEMENT NO | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO |
|---|---|---|---|
| N/A | N/A | N/A | N/A |

11. TITLE (Include Security Classification)

Rationale for a "Many Maps" Phonology Machine

12. PERSONAL AUTHOR(S)
Touretzky, David S., and Wheeler, Deirdre W.

| 13a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Technical | FROM 86Sept15 TO 91Sept14 | 1990, March | 10 |

16. SUPPLEMENTARY NOTATION
CMU-CS- 90-113 'Rationale for a 'Many Maps' Phonology Machine" will appear in the Proceeding of EMCSR-90: the 10th Europ. Mtg on Cybernetics Systems Research, Vienna, Austria, April 1990 World Scientific Publishing Co.

| 17. COSATI CODES | | |
|---|---|---|
| FIELD | GROUP | SUB-GROUP |
| | | |
| | | |

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Linguistics; phonology; connectionist modeling; sequence manipulation.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

SEE REVERSE SIDE

| | |
|---|---|
| 20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION |
| 22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Alan L. Meyrowitz | 22b. TELEPHONE (Include Area Code) (202) 696-4302; 22c. OFFICE SYMBOL N00014 |

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

## Abstract

$M^3P$, our "many maps" model of phonology, raises a number of questions about the nature of linguistic explanations and the ways in which connectionist models can contribute to the advancement of phonological theory. In this paper we attempt to answer some of the questions we and others have raised as a result of this work. We consider four sources of possible phonological constraints, and argue that articulatory and intelligibility constraints are insufficient to fully account for human phonological behavior. Computational constraints such as those suggested by our connectionist model may provide a solution.

# Rationale for a "Many Maps" Phonology Machine

DAVID S. TOURETZKY
School of Computer Science, Carnegie Mellon University
Pittsburgh, Pennsylvania, 15213, USA

DEIRDRE W. WHEELER
Department of Linguistics, University of Pittsburgh
Pittsburgh, Pennsylvania, 15260, USA

**Abstract**

M$^3$P, our "many maps" model of phonology, raises a number of questions about the nature of linguistic explanations and the ways in which connectionist models can contribute to the advancement of phonological theory. In this paper we attempt to answer some of the questions we and others have raised as a result of this work. We consider four sources of possible phonological constraints, and argue that articulatory and intelligibility constraints are insufficient to fully account for human phonological behavior. Computational constraints such as those suggested by our connectionist model may provide a solution.

## 1. Introduction

For a variety of reasons, we view phonology as an attractive starting point for cognitive scientists seeking to understand language. The domain, sequences of phonemes, is purely symbolic. The operations are familiar: chiefly insertion, deletion, and mutation of elements. The structures involved are quasi-linear. (Some theories employ limited-depth trees to represent syllables or feature hierarchies, but phonology does not admit self-similar embedded structures or objects of unbounded depth of the sort required by syntax.) True phonological processes are highly regular: they do not suffer from the plethora of special cases that complicates syntax and morphology. Even in those processes that are morphologically conditioned (and thus not purely phonological), such as the English /k/→/s/ rule that derives "electricity" from "electric" plus "-ity," the phoneme transformation itself is regular. Complexity comes only from the attachment of a morphological condition to the rule's environment.

Phonology is a unique language component for yet another reason: it is an autonomous process, not intertwined with higher levels in the way that syntax and semantics mutually interact. We acknowledge the existence of morphophonemic processes, but even here, influence flows in only one direction: morphological conditioning of phonological rules. Phonological processes do not interact with morphology. Autonomy is perhaps the most compelling reason why we see phonology as an approachable

domain for connectionist modeling. It is a domain where one may hope to achieve decisive and convincing results about human capabilities by pursuing a computation-driven approach.

Our goal in developing $M^3P$, our "Many Maps" Model of Phonology, is not to merely imitate human behavior or to implement a pre-existing theory. Rather, it is to investigate the ways in which adopting a particular model of computation—and its accompanying biologically-inspired constraints—can direct the development of linguistic theories, and even provide motivations for constraints on linguistic processes.

## 2. Overview of the Model

$M^3P$ began as an attempt to implement George Lakoff's theory of cognitive phonology (Lakoff, 1988; Lakoff, 1989) in connectionist hardware. An early, incomplete version was described in (Touretzky, 1989). The current version is described in (Wheeler & Touretzky, 1989; Touretzky & Wheeler, 1989; Touretzky & Wheeler, in press). This version differs substantially from Lakoff's proposal—a reflection of the theoretical progress made as the model has matured.

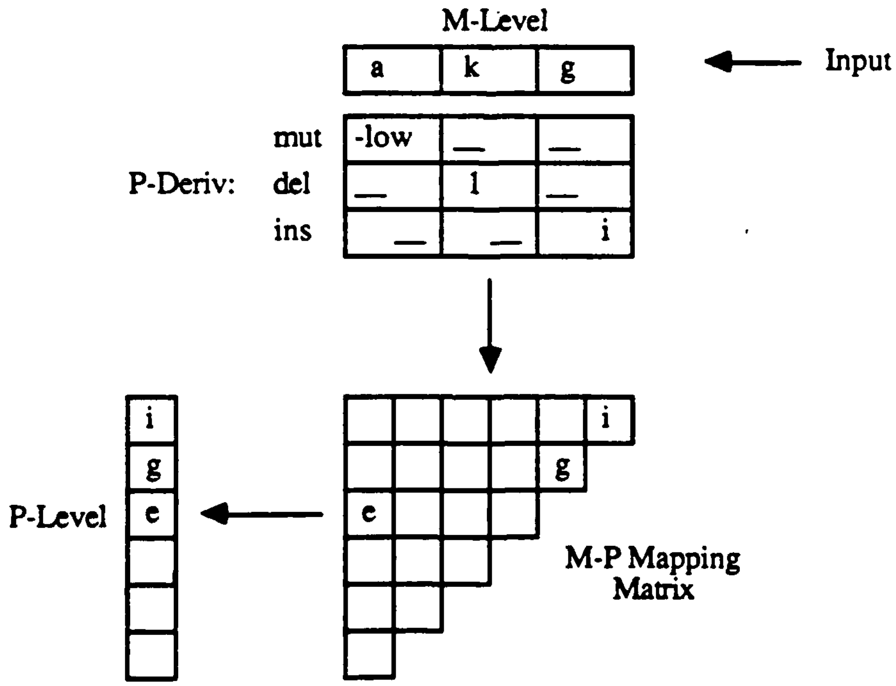


Figure 1: Example of the model's M-P map.

We cannot fully describe the details of the model here, but Figure 1 gives the flavor of our approach. This figure shows how an M (morpho-phonemic) level representation

of an example utterance, /akg/, is mapped to a P (phonemic) level representation, /egi/. In this artificial example the derivation involves three changes to the string: a mutation, a deletion, and an insertion. The changes are described in a "change buffer" called P-deriv. The M-level and P-deriv buffers both feed into an M-P mapping matrix whose job is to derive the phonemic representation, right-justified in the P-level output buffer, in one parallel step. The mapping matrix assures that there are no gaps or collisions in the output caused by multiple simultaneous insertions and deletions.

How do changes get written into the change buffer? One way is via M-P constructions (the counterpart of "rules" in traditional generative phonology) which examine the M-level representation and insert changes into P-deriv. Our model also contains a clustering mechanism that makes it possible to recognize clusters of adjacent segments sharing some property. Clustering provides an alternative to the traditional iterative accounts of phenomena such as vowel harmony, or voicing assimilation in consonant clusters, reminiscent of autosegmental representations. The M-level cluster modules are implemented using additional maps; cluster constructions read the states of these modules and write their changes into P-deriv, just as the M-P constructions do.

Most recently we have added a syllabifier to our model. Many insertion and deletion phenomena can be explained by the requirement that utterances be organized into well-formed syllables. Our syllabifier provides additional input into the mapping matrix, so that, for example, unsyllabified M-level segments will not appear at P-level, in effect causing the segment to be deleted.

The P-level representation is then fed through a second, very similar bit of mapping hardware called the the P-F map (not shown in Figure 1) to derive the F-level (phonetic) representation of the utterance. This P-F mapping is controlled by P-F constructions plus the special constructions associated with P-level clustering modules.

Our model is not capable of arbitrary string transformations. Its behavior is tightly constrained by a combination of factors: the mapping matrix can only perform insertions, deletions, and mutations of segments; there are only two levels of derivation, M-P and P-F; the clustering modules are highly specialized, and just powerful enough to model actual phonological phenomena such as vowel harmony; the syllabification mechanism is equally specialized. The key question that remains is: what is the relationship between the constraints on the model's behavior and the constraints that human beings appear to observe?

## 3. Sources of Phonological Constraints

In trying to account for the nature of of human phonological behavior, we find there are four sources of potential constraints. We consider them in turn.

### 3.1. Articulation

The first source of constraints is articulatory convenience: some sounds are simply easier to make than others. For example, affricates (the *ch* sound in "church") are relatively rare—English being an exception—while alvelar fricatives such as /s/ occur in many languages. Likewise, the basic vowels /a,i,u/ occur almost universally, while ü, ʌ̃ and ɔ are less common. Another example is consonant clusters. Many languages severely limit the ways in which consonants may adjoin; some have a strict CV structure in which consonants are always separated by vowels. (We note that insertion rules are frequently motivated by a syllabifier's desire to break up "unprorounceable" consonant clusters.) In contrast, English permits a variety of tri-consonantal clusters, such as /spl/ ("splash") and /skr/ ("scrap"). Many languages permit even more complex clusters.

Articulatory constraints are only weak constraints, because sequences that speakers of one language find unpronounceable may sound natural to other communities. However, it is still possible to objectively classify certain sounds or sound sequences as more marked than others, based on articulatory effort. If one looks at a wide spectrum of languages, the more marked sequences appear less frequently.

### 3.2. Computability

The second source of potential constraints on phonology is computability. Computational constraints reflect fundamental limitations imposed by the wiring of the brain's language production areas. These are the hard, universal constraints we are attempting to capture in $M^3P$. By its very nature, the model is unable to perform certain types of string transformations, such as reversing the order of phonemes in an utterance, or permuting the first and last consonant of a word. The model therefore predicts that no human language could possibly do these things.

We see a close relationship between our approach to phonology and the notion of parameters in (Chomsky, 1988). Chomsky suggests that languages can be characterized by particular sets of parameter values. The number of these parameters, their meanings, and their range of legal values is the province of Universal Grammar. The job of the language learner is to determine the particular parameter settings in use in his or her linguistic environment. This proposal has been put into practice by Dresher and Kaye (1990), who describe a mechanism for learning a language's metrical structure from examples. (The Dresher and Kaye model was first brought to our attention by Eric Nyberg, who suggests an alternative approach to the parameter setting problem in (Nyberg, 1989).)

The notion of a genetically-determined language machine configured by parameter values is certainly in harmony with our $M^3P$ model. However, language universals can presumably be captured by more than one parameter scheme (and more than one machine architecture), just as there are several distinctive feature systems that adequately

characterize phonemes. Chomsky leaves open the questions of where parameter systems come from and how to choose among them when evaluating theories of universal grammar. Our work provides a way to compare alternative parameter systems by looking at the underlying machine architecture each assumes. Criteria include, for example, circuit complexity, circuit depth, and degree of parallelism.

### 3.3. Intelligibility

Geoffrey Hinton (personal communication) has suggested a third source of phonological constraints: intelligibility. The hearer must be able to decode the speech signal. This limits the types of transformations speakers may make. Certain types of phonological transformations might simply be too difficult for the hearer to invert to arrive back at the correct underlying form.

There is a tension between the intelligibility and computability constraints. Both can account for the fact that speakers don't invert the phoneme order of entire words. But the intelligibility constraint does not seem powerful enough to explain all the peculiarities of human phonology. For example, switching the first and last consonant of a word wouldn't seem to interfere too much with intelligibility. Also, neutralization processes such as vowel reduction in English or devoicing of word final stops in German actually work against intelligibility in favor of articulatory convenience. Therefore, intelligibility seems too weak a constraint to fully account for the structure of phonology.

On the other hand, the computational constraints we have been proposing could turn out to be too powerful. People are remarkably flexible; it would be difficult to conclusively demonstrate that they are fundamentally incapable of certain types of phonological behavior. One objection frequently raised in response to $M^3P$'s computational constraints is Pig latin, a language game in which, for example, "games are fun" becomes "ames-gay are-way un-fay." This involves the sort of movement operation which our theory predicts is phonologically impossible. However, we doubt that competent Pig latin speakers are employing their natural phonological machinery to accomplish this task. Unlike a real language, Pig latin requires the speaker to first produce the correct English *surface* form of each word, and then transform it by one of two simple rules depending on whether the word begins with a vowel. The automaticity of Pig latin speech (after sufficient practice) does not imply that Pig latin rules have become part of the speaker's phonology. Many automatic processes, such as syntax, or the ability to play the piano, are non-phonological. What counts as "phonology" in our view are those processes that occur naturally in human languages and can be acquired automatically and *unconsciously* by children in a suitable linguistic environment. In contrast, Pig latin speakers must acquire a conscious representation of the rules of the game before they can speak it correctly.

## 3.4. History

A fourth source of potential phonological constraints, suggested to us by Prahlad Gupta (personal communication), is history. There is no way to know if the 5000 or so languages that have developed during the course of human history fully exercise our linguistic abilities. We must be careful to avoid turning historical accidents into universal principles. Just because a particular process has never occurred historically, this does not imply *a fortiori* that it is constrained from occurring some time in the future. Thus, one would be wise not to formulate overly-specific constraints based solely on historical evidence. The computational approach is helpful by providing additional motivation for certain types of constraints, but grey areas do remain. For example, to the best of our knowledge, no human language utilizes a quality-sensitive stress rule, such as "stress the penultimate high vowel in the word." This could be a legitimate computational constraint, but it might also be a historical artifact.

# 4. Relationship to Neuroscience

What is known to date about the neural basis of language comes largely from clinical studies (and subsequent autopsies) of stroke and head injury patients. Recently, some interesting new results have been obtained with radioactive imaging techniques that map metabolic activity throughout the brain during performance of language-related tasks. Along with the autopsy data, this is another valuable source of clues about how cognitive functions are distributed across different cortical regions. However, at this time there is still no detailed theory of how linguistic information is physically represented and processed in the brain.

$M^3P$ should not be taken literally as a biological model. We don't expect to find such clean representations and regular wiring structures in real brain tissue. Furthermore, our model does not yet account for developmental processes, speech errors, or the various types of aphasias people exhibit.

We do not wish to suggest, though, that our model is completely divorced from the neural level, the way Chomsky's theories are divorced from actual computation. We are making strong claims about the functional nature of the brain's language areas. Specifically, we predict that certain types of phonological processes are impossible, because they are incompatible with the $M^3P$ architecture. Even if the wiring of our model differs significantly from real neural circuitry, as we know it must, we assert that at the functional level the two systems may be equivalent.

Our work can be usefully contrasted with backpropagation-based approaches to phonology, such as the Rumelhart and McClelland (1986) verb learning model, or the more recent work using sequential recurrent nets by Gasser and Lee (1989). These models can in principle learn any input-output mapping, so they are unable to impose hard

constraints on phonological operations the way $M^3P$ does. In fact, they don't appear to offer any explanation for the extremely regular and symbolic nature of phonology. At best, they may correctly mimic human phonological behavior. Often, though, their behavior is not completely correct after training.

Since rules have no independent existence in a backprop-based model other than as facets of a monolithic input-output mapping, each combination of rule interactions must be learned as a separate case. There is no mechanism requiring the individual rules to interact systematically. Backprop-based systems without internal structure, if they are powerful enough to model real human phonology, can just as easily learn many nonphonological behaviors. Unlike the Rumelhart and McClelland model, Gasser and Lee's sequential network cannot easily learn reversals and long-range metatheses, but this limitation also hinders it from modeling *any* process that examines segments from right to left, such as regressive feature spreading or assignment of stress to penultimate syllables, both of which are common in human languages.

Even if a more sophisticated backprop-based model could be constructed whose generalization behavior were completely correct, an important question would remain. Is the brain fundamentally incapable of certain types of phonological operations, as $M^3P$ predicts, or is its phonological machinery as unrestricted as the backprop models, merely adapting itself in response to linguistic inputs? This is, of course, an instance of the classic *rationalism vs. empiricism* controversy. We are content to toil in the rationalist camp, enhancing our model and predicting constraints that linguists may seek to verify or refute. Some day, neurolinguistics may provide the decisive answer to the rationalist/empiricist debate. If so, we are confident that such progress will have been made possible, at least in part, by the work of connectionist modelers in both camps.

**Acknowledgements**

This research was supported by a contract from Hughes Research Laboratories, by the Office of Naval Research under contract number N00014-86-K-0678, and by National Science Foundation grant EET-8716324. We thank Gillette Elvgren III for his work on implementing the simulations, and George Lakoff, Geoffrey Hinton, and Prahlad Gupta for stimulating conversation.

## References


[1] Chomsky, N. (1988) Language and Problems of Knowledge. Cambridge, MA: MIT Press.

[2] Dresher, B. E., and Kaye, J. D. (1990) A computational learning model for metrical phonology. To appear in Cognition.

[3] Gasser, M., and Lee, C. (1989) Networks that learn phonology. Computer Science technical report number 300, Indiana University.

[4] Lakoff, G. (1988) A suggestion for a linguistics with connectionist foundations. In D. S. Touretzky, G. E. Hinton, and T. J. Sejnowski (eds.), Proceedings of the 1988 Connectionist Models Summer School, pp. 301-314. San Mateo, California: Morgan Kaufmann.

[5] Lakoff, G. (1989) Cognitive phonology. Draft of paper presented at the UC-Berkeley Workshop on Constraints vs Rules, May 1989.

[6] Nyberg, E. H. (1989) Weight propagation and parameter setting. Ph.D. thesis proposal, computational linguistics program, Carnegie Mellon University.

[7] Touretzky, D. S. (1989) Toward a connectionist phonology: the "many maps" approach to sequence manipulation. Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, pp. 188-195. Hillsdale, NJ: Erlbaum.

[8] Touretzky, D. S., and Wheeler, D. W. (1989) Two derivations suffice: the role of syllabification in cognitive phonology. Manuscript.

[9] Touretzky, D. S., and Wheeler, D. W. (in press) A computational basis for phonology. In D. S. Touretzky (ed.), Advances in Neural Information Processing Systems 2. San Mateo, California: Morgan Kaufmann. To appear April, 1990.

[10] Wheeler, D. W., and Touretzky, D. S. (1989) A connectionist implementation of cognitive phonology. Technical Report CMU-CS-89-144, Carnegie Mellon University School of Computer Science. To appear in G. Lakoff and L. Hyman (eds.), Proceedings of the UC-Berkeley Phonology Workshop on Constraints vs, Rules, University of Chicago Press.